Year: 2020

# The mechanisms responsible for improved information transfer in avatar-based patient monitoring: A Multicenter Comparative Eye-Tracking Study

Tscholl, David Werner ; Rössler, Julian ; Handschin, Lucas ; Seifert, Burkhardt ; Spahn, Donat R ; Nöthiger, Christoph B

Abstract: Background: Patient monitoring is central to the safety of state-of-the-art perioperative and intensive care medicine. While current state-of-the-art patient monitors display vital signs in the form of numbers and curve forms, Visual Patient technology creates an easy to interpret virtual patient avatar model, which, in a previous study, enabled anesthesia providers to perceive more vital sign information during short glances than conventional monitoring. In this study, we used eye-tracking technology to study the deeper mechanisms underlying information perception in both, conventional and avatar-based patient monitoring. Objective:In this study, we used eye-tracking technology to study the deeper mechanisms underlying information perception in both, conventional and avatar-based patient monitoring. Methods: In this prospective, multi-center study with a within subject design, we showed 32 anesthesia experts (physicians and nurse anesthetists) a total of four 3- and 10-second monitoring scenarios alternatingly as either routine conventional or avatar-based monitoring in random sequence. All participants observed the same scenarios with both monitoring technologies. After each scenario, we asked participants to report the status of the vital signs. Using an eye-tracker, we recorded the participants' gaze paths as they observed the scenarios. From the eye-tracking recordings, we evaluated which vital signs the participants had visually fixated, how often and for how long during a scenario, and therefore, could have potentially read and perceived this vital sign. We compared the frequencies and durations with which the participants had visually fixated the vital signs between the two monitoring technologies. Results: Participants visually fixated more vital signs per scenario, median (IQR): 10 (9-11) vs. 6 (4-8), $p<0.001$ in avatar-based monitoring (median of differences: 3 vital signs (95% confidence interval [95%CI 3-4]). In a multivariable linear regression, monitoring technology (conventional vs. avatar-based monitoring, difference -3.3, $p<0.001$) was an independent predictor of the number of visually fixated vital signs. Only scenario duration affected the difference in vital sign fixations between technologies. The difference was more prominent in 3-second scenarios, difference -1.5, $p=0.04$. Study center, profession, gender, and scenario order did not influence the differences between conventional and avatar-based monitoring. In all four scenarios, the participants visually fixated nine of the 11 total vital signs shown statistically significantly longer using the avatar (all $p < 0.001$). Four critical vital signs, i.e., pulse rate, blood pressure, oxygen saturation, and respiratory rate were visible almost the entire time of a scenario with avatar-based monitoring, while with conventional monitoring, these were only visible for fractions of the observations. Visual fixation of a certain vital sign was associated with the correct perception of that certain vital sign in both technologies. Phi coefficient for avatar: 0.358, for conventional monitoring: 0.515, both $p<0.001$. Conclusions: This study uncovered, by use of eye-tracking, one of the mechanisms responsible for the improved information transfer in avatar-based monitoring. The design of the avatar technology, which presents the information about multiple vital signs integrated into forms and colors of the corresponding anatomical parts of a patient model enabled parallel perception of multiple vital signs, and thereby increased the number of visually fixated vital signs and the time available to fixate each vital sign. With this finding confirmed by eye-tracking, this study adds a new and higher level of empirical evidence as to why avatar-based monitoring improves the perception of vital sign information compared to conventional monitoring.

DOI: https://doi.org/10.2196/15070

# The mechanisms responsible for improved information transfer in avatar-based patient monitoring: A Multicenter Comparative Eye-Tracking Study.

David Werner Tscholl, Julian Rössler, Lucas Handschin, Burkhardt Seifert, Donat R. Spahn, Christoph Nöthiger

Submitted to: Journal of Medical Internet Research
on: June 17, 2019

# Table of Contents

# The mechanisms responsible for improved information transfer in avatar-based patient monitoring: A Multicenter Comparative Eye-Tracking Study.

David Werner Tscholl, MD; Julian Rössler; Lucas Handschin, MD; Burkhardt Seifert, PhD; Donat R. Spahn, MD; Christoph Nöthiger, MD

**Corresponding Author:**
David Werner Tscholl, MD
Phone: +41+41786636787
Email: davidtscholl@me.com

## *Abstract*

**Background:** Patient monitoring is central to the safety of state-of-the-art perioperative and intensive care medicine. While current state-of-the-art patient monitors display vital signs in the form of numbers and curve forms, Visual Patient technology creates an easy to interpret virtual patient avatar model, which, in a previous study, enabled anesthesia providers to perceive more vital sign information during short glances than conventional monitoring. In this study, we used eye-tracking technology to study the deeper mechanisms underlying information perception in both, conventional and avatar-based patient monitoring.

**Objective:** In this study, we used eye-tracking technology to study the deeper mechanisms underlying information perception in both, conventional and avatar-based patient monitoring.

**Methods:** In this prospective, multi-center study with a within subject design, we showed 32 anesthesia experts (physicians and nurse anesthetists) a total of four 3- and 10-second monitoring scenarios alternatingly as either routine conventional or avatar-based monitoring in random sequence. All participants observed the same scenarios with both monitoring technologies. After each scenario, we asked participants to report the status of the vital signs. Using an eye-tracker, we recorded the participants' gaze paths as they observed the scenarios. From the eye-tracking recordings, we evaluated which vital signs the participants had visually fixated, how often and for how long during a scenario, and therefore, could have potentially read and perceived this vital sign. We compared the frequencies and durations with which the participants had visually fixated the vital signs between the two monitoring technologies.

**Results:** Participants visually fixated more vital signs per scenario, median (IQR): 10 (9-11) vs. 6 (4-8), $p<0.001$ in avatar-based monitoring (median of differences: 3 vital signs (95% confidence interval [95%CI 3-4]). In a multivariable linear regression, monitoring technology (conventional vs. avatar-based monitoring, difference -3.3, $p<0.001$) was an independent predictor of the number of visually fixated vital signs. Only scenario duration affected the difference in vital sign fixations between technologies. The difference was more prominent in 3-second scenarios, difference -1.5, $p=0.04$. Study center, profession, gender, and scenario order did not influence the differences between conventional and avatar-based monitoring. In all four scenarios, the participants visually fixated nine of the 11 total vital signs shown statistically significantly longer using the avatar (all $p < 0.001$). Four critical vital signs, i.e., pulse rate, blood pressure, oxygen saturation, and respiratory rate were visible almost the entire time of a scenario with avatar-based monitoring, while with conventional monitoring, these were only visible for fractions of the observations. Visual fixation of a certain vital sign was associated with the correct perception of that certain vital sign in both technologies. Phi coefficient for avatar: 0.358, for conventional monitoring: 0.515, both $p<0.001$.

**Conclusions:** This study uncovered, by use of eye-tracking, one of the mechanisms responsible for the improved information transfer in avatar-based monitoring. The design of the avatar technology, which presents the information about multiple vital signs integrated into forms and colors of the corresponding anatomical parts of a patient model enabled parallel perception of multiple vital signs, and thereby increased the number of visually fixated vital signs and the time available to fixate each vital sign. With this finding confirmed by eye-tracking, this study adds a new and higher level of empirical evidence as to why avatar-based monitoring improves the perception of vital sign information compared to conventional monitoring.

(JMIR Preprints 17/06/2019:15070)
DOI: https://doi.org/10.2196/preprints.15070

### Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

# Original Manuscript

# *The mechanisms responsible for improved information transfer in avatar-based patient monitoring: A Multicenter Comparative Eye-Tracking Study.*


## Authors

David W. Tscholl[1] M.D., Julian Rössler[1], Lucas Handschin[1] M.D., Burkhardt Seifert[2] Ph.D., Donat R. Spahn[1] M.D. F.R.C.A., Christoph B. Noethiger[1] M.D.

## Affiliations/Institutional addresses

[1]Institute of Anesthesiology, University and University Hospital Zurich, Raemistrasse 100, 8091 Zurich, Switzerland

[2]Department of Biostatistics, Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Hirschengraben 84, 8001 Zurich, Switzerland.

## Corresponding author

David W. Tscholl, M.D.

Institute of Anesthesiology, University Hospital Zurich, Raemistrasse 100, 8091 Zurich, Switzerland

Phone: +41 (0)44 253 94 51, Fax: +41 (0)44 255 44 09

E-mail: David.Tscholl@usz.ch



## MESH Keywords

Computers; Diagnosis; Visual perception; Awareness; Patient Safety.


## Competing interests

The University of Zurich, Zurich Switzerland and Koninklijke Philips N.V., Amsterdam, Netherlands entered a joint development and licensing agreement to develop avatar-based monitoring software based on technology owned by the University and described in this manuscript. As part of their contract with the University, as designated inventors, the authors DWT and CBN could receive royalties.


## Funding

Institute of Anesthesiology, University, and University Hospital Zurich: institutional funding.

University of Zurich: Proof of concept funding (UZ16/288PoC).

University of Zurich: DWT Filling-the-gap career development grant.


## Authors' contributions

DWT, LH, BS, DRS, and CBN helped to design the study.

DWT, LH and CBN helped to collect the data.

DWT, JR, LH, BS, DRS, and CBN helped to analyze the data.

DWT, JR, LH, BS, DRS, and CBN helped to write the manuscript and approved the final version.

## Acknowledgements

The authors are thankful to the study participants for their time and effort.

## *List of abbreviations*

| | |
|---|---|
| NASA | National Aeronautics and Space Administration |
| TLX | Task Load Index |
| IQR | Interquartile range |
| MoD | Median of differences |
| 95%CI | 95% confidence interval |

# Abstract

## Background

Patient monitoring is central to the safety of state-of-the-art perioperative and intensive care medicine. While current state-of-the-art patient monitors display vital signs in the form of numbers and curve forms, Visual Patient technology creates an easy to interpret virtual patient avatar model, which, in a previous study, enabled anesthesia providers to perceive more vital sign information during short glances than conventional monitoring. In this study, we used eye-tracking technology to study the deeper mechanisms underlying information perception in both, conventional and avatar-based patient monitoring.

*Methods*

In this prospective, multi-center study with a within subject design, we showed 32 anesthesia experts (physicians and nurse anesthetists) a total of four 3- and 10-second monitoring scenarios alternatingly as either routine conventional or avatar-based monitoring in random sequence. All participants observed the same scenarios with both monitoring technologies. After each scenario, we asked participants to report the status of the vital signs. Using an eye-tracker, we recorded the participants' gaze paths as they observed the scenarios. From the eye-tracking recordings, we evaluated which vital signs the participants had visually fixated, how often and for how long during a scenario, and therefore, could have potentially read and perceived this vital sign. We compared the frequencies and durations with which the participants had visually fixated the vital signs between the two monitoring technologies.

*Results*

Participants visually fixated more vital signs per scenario, median (IQR): 10 (9-11) vs. 6 (4-8), $p<0.001$ in avatar-based monitoring (median of differences: 3 vital signs (95% confidence interval [95%CI 3-4]). In a multivariable linear regression, monitoring technology (conventional vs. avatar-based monitoring, difference -3.3, $p<0.001$) was an independent predictor of the number of visually fixated vital signs. Only scenario duration affected the difference in vital sign fixations between technologies. The difference was more prominent in 3-second scenarios, difference -1.5, $p=0.04$. Study center, profession, gender, and scenario order did not influence the differences between conventional and avatar-based monitoring. In all four scenarios, the participants visually fixated nine of the 11 total vital signs shown statistically significantly longer using the avatar (all $p < 0.001$). Four critical vital signs, i.e., pulse rate, blood pressure, oxygen saturation, and respiratory rate were visible almost the entire time of a scenario with avatar-based monitoring, while with conventional monitoring, these were only visible for fractions of the observations. Visual fixation of a certain vital sign was associated with the correct perception of that certain vital sign in both technologies. Phi coefficient for avatar: 0.358, for conventional monitoring: 0.515, both $p<0.001$.

*Conclusion*

This study uncovered, by use of eye-tracking, one of the mechanisms responsible for the improved information transfer in avatar-based monitoring. The design of the avatar technology, which presents the information about multiple vital signs integrated into forms and colors of the corresponding anatomical parts of a patient model enabled parallel perception of multiple vital signs, and thereby increased the number of visually fixated vital signs and the time available to fixate each vital sign. With this finding confirmed by eye-tracking, this study adds a new and higher level of empirical evidence as to why avatar-based monitoring improves the perception of vital sign information compared to conventional monitoring.

## Introduction

The World Health Organization considers continuous patient monitoring to be "extremely important" for the safety of the more than 313 million patients worldwide undergoing surgery each year.[1, 2] In operating rooms and intensive care units around the world, monitors help millions of health care providers every day to make critical treatment decisions.[3, 4] However, previous research has found that conventional patient monitoring based on numbers and waveforms is not ideally suited for transferring patient status information to health care providers. Authors recommended the development of new technologies to improve information transfer, especially from short glances at the monitors, because that's how studies found care providers to perform monitoring in real-life.[5-7]

In a previous comparative study with conventional monitoring, we found that anesthesia professionals were able to perceive more vital signs when monitoring with Visual Patient, a technology integrating the vital sign information into an easy to interpret animated avatar model of the patient's situation, designed according to principles of user-centered design.[4, 8, 9] Also, using avatar-based monitoring, the participants rated their self-confidence in the correctness of their diagnoses as higher, and their subjectively perceived workload as lower.

While the bi-ocular human visual field encompasses approximately 214 arc degrees horizontally and 150 arc degrees vertically, we can only see sharply in a circular area of about two arc degrees in the center of our visual field, named the fovea.[10] During reading we move our eyes to let the light reflected from the words, or the numbers fall through the pupil and the lens directly onto the fovea. At a distance of about one arm's length, the foveal region in which we can see sharp, colorful images has approximately the size of a thumbnail or a circle with a radius of two centimeters.[11, 12]

Human eye movements take place in the form of so-called fixations and saccades. Visual fixations are the periods during which the gaze rests on a position, and information can reach the visual cortex and potentially be interpreted. Saccades are the rapid movements of the eyeballs between fixations.[13] Eye-tracking technology can record both visual fixations and saccades. For this study, we systematically recorded, analyzed and compared eye-tracking data of participants who watched patient monitoring scenarios alternatingly as conventional and avatar-based patient monitoring. The rationale of this study was to uncover underlying functional principles in both monitoring technologies through eye-tracking analysis. These results may be useful for improved understanding of avatar-based monitoring but also, across domains, for the future development of user interfaces designed to transfer relevant information as efficiently as possible. We hypothesize that the Visual Patient facilitates information perception through its compact layout, which enables users to visually fixate more vital signs.

## Methods

In this paper, we describe the analysis of eye-tracking data that was collected as part of a multi-method laboratory study. The primary objective of that study was to compare the perceptual performance of anesthesia professionals using a newly developed avatar-based technology with state-of-the-art number and wave-form-based patient monitoring.[8]

The Ethics Committee of the Canton of Zurich, Switzerland, reviewed the study protocol and issued a declaration of non-responsibility clarifying that the research project does not fall into the scope of the Human Research Act (Business Administration System for Ethics Committees, number: 2016-00103). Nevertheless, we obtained written consent to the use of the collected data for scientific purposes from all participants.

### Description of Visual Patient Technology

The Visual Patient Technology, as used in this study, can display the 11 most frequently monitored vital signs: Pulse rate, blood pressure, oxygen saturation, ST-segment of the electrocardiogram, central venous pressure, respiratory rate, tidal volume, expiratory carbon dioxide concentration, body temperature, brain activity and degree of neuromuscular relaxation. We developed the technology as a situation awareness tool in analogy to the Synthetic Vision Technology in aviation, according to the principles of user-centered design and principles of logic.[4,14] Synthetic Vision technology generates a virtual image of the environment from the data measured by the aircraft, e.g., airspeed, and Global Positioning System geolocation data and data stored on onboard computers, e.g., geo-referenced terrain elevation data. For the pilot, the generated virtual image looks like the view from the cockpit in perfect weather conditions. This similarity between the virtual image and reality makes the image intuitively understandable and thus enables a quick and uncomplicated perception of the flight situation. Visual Patient Technology uses the same logic by creating a virtual image of the patient from vital data. It presents the data in a way that corresponds to the real phenomena that the data would cause in the patient. For example, high brain activity is represented by open eyes, because that is what the care providers expect from a patient with high brain activity according to their mental models.

This so-called direct presentation of information eliminates the need to calculate the relevant information, e.g., "What is the current anesthesia depth?" from lower-level data, e.g., bispectral Index = 85.[15] Besides this direct presentation of information, the other two main features of the avatar technology are the pre-processing of data for each vital sign into the categories "no data measured", "too low", "normal" or "too high" and the presentation of vital parameter information in multiple visualizations simultaneously. For example, the caregiver can evaluate the respiratory rate based on the respiratory rate of the avatar lung and the formation rate of the carbon dioxide cloud exhaled by the avatar.

These combined functions translate a large number of numerical values into an animated model of the patient situation, which the caregiver can evaluate and save at a glance. The vital parameters are translated into the avatar model in real time from the monitoring data. We have described the validation and evaluation process of the avatar in detail in previous studies.[8,9,21]

## Limitations of Visual Patient Technology

Visual patient technology has inherent limitations that we would like to mention. The technology simplifies vital signs into categories ( i.e., "too low", "normal", or "too high"). This pre-processing leads to improved intelligibility and diagnostic certainty but also reduces data accuracy (3 discrete categories vs. 300 different numbers in the case of pulse rate). Another limitation of the Visual Patient version used in this study is that it cannot yet display trends. This aspect is important because trend displays of conventional patient monitors can help care providers to detect changes over time. In this context, it is important to note that we are developing avatar-based monitoring to improve information transfer but not as a replacement of the conventional monitoring streams. Successful integration of the two technologies will be key for the success of the Visual Patient, as with synthetic vision technology and numerical flight data.

### Study participants

The participants of this study were attending and resident physician anesthesiologists and specialist anesthesia nurses from the anesthesia departments of the University Hospital Zurich and the Cantonal Hospital of Winterthur. The University Hospital Zurich is one of the largest University hospitals in Switzerland performing more than 30,000 operations per year, and the Cantonal Hospital Winterthur is a large regional teaching hospital with approximately 10,000 operations per year.

Participation in this study was voluntary and there was no monetary compensation for the participants. We recruited colleagues who responded to an institutional invitation and recruited additional colleagues according to availability. At both centers, we included equal numbers of male and female participants and participants from these three professional groups: 1. senior anesthesia physicians, 2. resident physicians, and 3. anesthesia nurses.

### Study setting

Before the beginning of the data collection, the participants received training in avatar-based monitoring by means of a 6-minute educational video (Supplementary video 2). The participants also familiarized themselves with the layout of the conventional monitoring used in the study: a simulation of a GE Datex Ohmeda Monitors (General Electric Company, Boston, MA, USA) recorded with the SimMon App (Castle 2 Andersen ApS, Hillerød, Denmark), which was equivalent to the patient monitoring in routine use in the two centers. There was no additional training in conventional patient monitoring because all participants had at least one full year, some even decades, of anesthesia experience. The eye-tracking data was recorded during the evaluation of patient monitoring scenarios. In random order, we presented participants with 3- and 10-second, prerecorded videos of patient monitoring scenarios shown in avatar and state-of-the-art number and wave-form format. Supplementary Video 2 provides examples of a conventional and an avatar-based scenario. Each participant rated four videos in sequence. These videos consisted of one 3- and one 10-second monitoring scenario, each of which was shown twice, once with either technology. The scenarios came from a pool of four total scenarios, as outlined in Supplementary Table 1. The scenarios were designed with unambiguously safe or unsafe vital sign values and contained random vital sign abnormalities, to avoid pattern recognition, i.e., inferring the status of vital signs based on the status of the other vital signs.

To blind the participants to the fact that they were evaluating the same scenarios twice (once with either technology), we showed the scenarios in alternating order, starting with a random first scenario. Supplementary Figure 1 shows a flowchart detailing this procedure. The scenario playback was performed on an Aspire V15 Nitro 15-inch laptop computer (ACER, Inc., Taipei, Taiwan) in ultra-high resolution (3840x2160 pixels) at 60 frames per second. The conventional monitoring scenarios included a standard audio signal with frequency and pitch for heart rate and oxygen saturation.

After brief time intervals, the screens darkened, and the participants indicated how they had perceived the 11 total vital signs displayed in the scenarios as either “too low”, “too high”, “safe” or “no recall.” We based this method on the Situation Awareness Global Assessment Tool developed by Endsley.[4, 16] Then, after each scenario, the participants, for every vital sign, indicated how confident they felt that they had perceived it correctly. Furthermore, they were asked to rate their subjectively perceived workload for each scenario using the National Aeronautics and Space Administration (NASA) Task Load Index (TLX).[17, 18] Data was collected using an iPad-based (Apple Inc., Cupertino, CA, USA) data collection tool.[19]

### *Recording and analysis of Eye-Tracking Data*

We evaluated the eye-tracking data for this study according to physiological principles of the human eye and neurophysiological principles of human vision outlined in the introduction. We used a stationary eye tracker (Gazepoint GP3 by Gazept, Vancouver, Canada) to capture visual fixations and saccades of participants observing conventional and avatar-based patient monitoring scenarios. The eye-tracker recorded the position of the foveal vision on the screen 60 times per second and with 0.5 - 1 degree of visual angle accuracy.

### *Outcome measures*

### Vital signs fixated per scenario (primary outcome)

We chose to compare visual fixations in this study because we regarded them as a relevant requirement for perception. Based on the anatomic and physiologic principles outlined in the introduction, we analyzed each visual fixation longer than 50 milliseconds in the eye-tracking recordings of each participant and scenario, and, for each visual fixation, counted those vital signs that lay within 2 centimeters of the fixation. Using this method, we identified the vital signs that participants could potentially have read during the recording because they lay within the potentially readable visual area. A video demonstrating this method is available in Supplementary Video 1. Information can only reach the brain for processing after prior reading, which in turn requires a visual fixation.

Conventional patient monitoring shows the vital signs on the screen in the form of numbers or waveforms. In conventional monitoring, if a participant had a visual fixation within 2 centimeters of the number or the waveform representing a certain vital sign, e.g., pulse rate, we counted a visual fixation for this certain vital sign.

Avatar-based monitoring visualizes the vital sign information in large, colorful visualizations in an animated patient model. These visualizations were designed to display a meaningful commonality with the reality they mirror. For example, in the avatar, the pulse rate corresponds to the pulsation of the avatar's body to represent the pulse wave passing through the body with every heartbeat. Accordingly, if a participant had a visual fixation within 2 centimeters of a visualization of the avatar representing a certain vital sign, e.g., the avatar's pulsating body, we counted a visual fixation for the respective vital sign. By comparing the numbers of vital sign visual fixations between the two technologies, we wanted to find out whether the participants could visually fixate more vital signs with either one of the two technologies.

## Visual fixations per vital sign

We also compared visual fixations for each of the 11 vital signs individually. This allowed us to determine whether vital signs were visually fixated more often with either one of the two monitoring technologies. We expected these findings may provide an explanation for the improved perceptive performance in avatar-based patient monitoring.

## Duration of visual fixations per vital sign

In analogy to the visual fixations per vital sign, we also compared the time durations of the visual fixations of each of the 11 vital signs with both monitoring technologies. In doing so, we evaluated whether either one of the two monitoring technologies would cause the vital signs to be visible for a longer time per observation. We analyzed this outcome measure because longer availability of the vital sign information could explain why participants' perceptual performance was improved with avatar-based patient monitoring.

## Correlation of vital sign visual fixations with correct perception

To evaluate the association of visual fixation of a vital sign and its correct perception, we calculated Phi correlation coefficients. If indeed visual fixation correlated with correct perception, and the avatar enabled more vital signs to be seen per time interval, these results may serve to validate both, the study method and the avatar concept.

## Correlation of vital sign visual fixations with perceived confidence

We calculated coefficients between visual fixation and diagnostic confidence to evaluate whether the visual fixation of a vital sign correlated with the subjectively perceived confidence in the correctness of one's own diagnosis.

### *Statistical analysis*

## Sample size calculation

Before starting the study, we conducted a pilot study with 5 participants. We calculated the sample size using the effect size of 1.23 measured in the pilot study. Assuming a clinically relevant difference of 1 vital sign and an observed standard deviation of 0.81, the posthoc power analysis for a paired t-test resulted in a sample size of 8 participants, for an alpha-error probability of 5% and a power of 0.8. To achieve this sample size in both centers and all 4 scenarios, we had to include at least 32 participants.

## Descriptive statistics and normality-tests

Distribution of variables is expressed using medians and interquartile ranges (IQR) regardless of normality. Normality was assessed with the Shapiro-Wilks test and visual inspection of quantile-quantile plots of dependent variables.

## T tests and nonparametric tests

Since participants watched and evaluated the same monitoring scenarios with both monitoring technologies, we were able to perform intra-participant comparisons and, depending on normality, used either paired Student t-tests or Wilcoxon signed rank test to compare the number of vital sign visual fixations with both monitoring technologies. We calculated the 95% confidence (95%CI) interval of the median of differences (MoD) using the Hodges-Lehmann estimate. To test the differences in visual fixations and duration of visual fixations per vital sign for statistical significance, we used Mann-Whitney tests. In this study, we performed multiple comparisons and, therefore, considered p-values between 0.05 and 0.01 as trends and p-values of <0.01 as statistically significant.

## Multivariable linear regression

Multivariable linear regression was performed with number of visually fixated vital signs between the monitoring technologies and its differences as dependent variables. Scenario duration, order of scenarios, center, profession and gender of the participant served as possible predictors. Clustering of observations within the same participant was addressed using cluster robust standard errors.

## Correlation analyses

To test for associations between visual fixation of a vital sign and its correct perception as well as participants' subjectively perceived confidence in the correctness of the diagnosis, we calculated chi-square tests for association and Pearson's Phi coefficients between visual fixation, accurate perception, and diagnostic certainty. The Phi coefficient corresponds to a Pearson correlation coefficient estimated for two binary variables. We considered "very unconfident" and "unconfident" as 0 and "confident" and "very confident" as 1. If the frequency of an event was lower than five, we used Fisher's exact test to assess statistical significance.

## Statistical software

We used Q*Power 3 (Heinrich-Heine-University, Düsseldorf, Germany),[20] Prism 8.1.1. (GraphPad

Software, La Jolla, CA, USA), IBM SPSS Statistics 24 (International Business Machines Corporation, Armonk, NY, USA), and Stata 13.1 (StataCorp, College Station, TX, USA).

# Results

*Study and participant characteristics*

Table 1 shows the characteristics of the study and participants in detail. A total of 32 participants participated in the two study centers. Had we been able to record data from all 32 participants and scenarios, a theoretical maximum of 64 direct comparisons between avatar-based and conventional patient monitoring would have been possible. However, we were unable to record eye-tracking data in two participants. In four more participants, we were only able to record one of the two monitoring scenarios they watched due to technical problems.

Despite the missing data, in this study, we were still able to evaluate 56 within subject comparisons of eye-tracking data (= 88% of all 64 theoretically possible comparisons).

**Table 1:** The study and participant characteristics. IQR=Interquartile range. GE=General Electric Company, Boston, MA, USA.

| Study center | University Hospital Zurich (N=16) | Cantonal Hospital Winterthur (N=16) | Total (N=32) |
|---|---|---|---|
| Total number of participants with successful eye-tracking recording | 16 | 14 | 30 |
| Number of direct comparisons | 32 | 24 | 56 |
| Number of staff members (%) | 6 (37) | 6 (37) | 12 (37) |
| Number of residents (%) | 4 (25) | 4 (25) | 8 (25) |
| Number of nurse anesthetists (%) | 6 (37) | 6 (37) | 12 (37) |
| Number of female/male participants (%) | 7 (44) / 9 (56) | 10 (62) / 6 (37) | 17 (53) / 15 (47) |
| Participants age groups (%): | | | |
| 25 to 34 years | 10 (63) | 6 (38) | 16 (50) |
| 35 to 44 years | 6 (38) | 2 (13) | 8 (25) |
| 45 to 54 years | 0 (0) | 6 (38) | 6 (19) |
| 55 to 65 years | 0 (0) | 2 (13) | 2 (6) |
| Participants anesthesia experience groups (%): | | | |
| Less than 1 year | 1 (6) | 1 (6) | 2 (6) |
| 1 to 5 years | 5 (31) | 4 (25) | 9 (28) |
| 5 to 10 years | 9 (56) | 1 (6) | 10 (31) |
| More than 10 years | 1 (6) | 10 (63) | 11 (34) |
| Median (IQR) number of monitors from different manufacturers previously used | 2 (2 - 3) | 2 (1 - 4) | 2 (1 - 3) |
| Median (IQR) duration of data collection session in minutes | 32 (28-35) | 35 (32-41) | 33 (30 - 39) |
| Duration of study in days | 20 | 2 | 22 |

## *Outcome measures*

### Vital signs fixated per scenario

All participants, in all scenarios, with avatar-based monitoring were able to visually fixate more vital signs than with conventional monitoring (Figure 1). In the short 3-second scenarios, the median numbers of vital sign fixations with avatar-based monitoring were about twice as high than with conventional patient monitoring. In scenario 1, avatar-based: 9 (interquartile range [IQR] 9-10) vs. conventional: 4 (IQR 4-6), p<0.001, Median of Differences (MoD): 3 (95%CI 3-4). In scenario 2: 9 (IQR 8-10) vs.: 5 (IQR 3-6), p<0.001, MoD: 5 (95%CI 2-6). In scenario 3, the first of the longer 10-second scenarios, the median number of vital sign fixations avatar-based was: 11 (IQR 11 -11) vs. conventional: 9 (IQR 6-10), p=0.002, MoD: 2 (95%CI 0-4). In scenario 4, the second 10-second scenario, vital sign visual fixations were: 11 (IQR 11-11) vs. 8 (IQR 7-10), p<0.001, MoD: 3 (95%CI 1-4). Figure 1 shows these results on individual participant level.

In a multivariable linear regression adjusted for scenario duration, order of scenarios, center, profession and gender of the participant, the technology (conventional vs. avatar-based monitoring) had a significant effect on the number of vital sign fixations: difference -3.28, 95%CI -3.86 to -2.69, p<0.001 (F[6, 30] = 145, Prob>F < 0.001, R-squared = 0.56, adjusted for clusters within participants). Table 2 shows the results of the multivariable linear regression for the difference of numbers of visually fixated vital signs between the technologies. In this analysis, only scenario duration affected the difference in vital sign fixations between technologies. The difference was more prominent in 3-second scenarios, difference -1.46, p=0.04. Study center, profession, gender, and scenario order did not influence the differences between conventional and avatar-based monitoring.

**Figure 1:** Avatar-based monitoring compared to conventional patient monitoring: Number of vital signs visually fixated on an individual participant level. N for scenario 1 (3sec) was 12, for scenario 2 (3sec) 15, for scenarios 3 (10sec) 14, and for scenario 4 (10sec) 15. The dotted lines indicate the medians. Participants 1-8 (University Hospital Zurich) and 17-24 (Cantonal Hospital Winterthur) rated scenarios 1 and 3, participants 9-16 (University Hospital Zurich) and 25-32 (Cantonal Hospital Winterthur) rated scenarios 2 and 4.

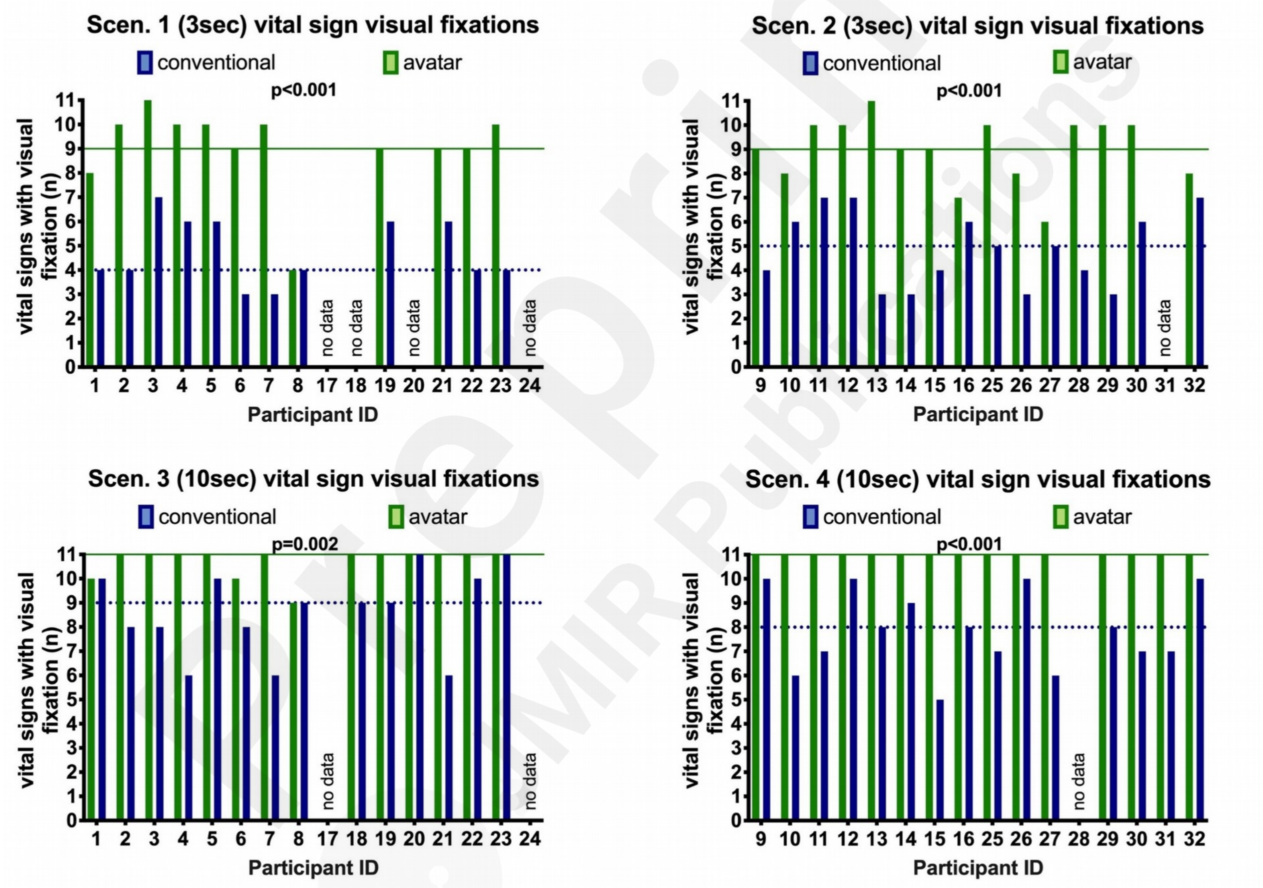


**Table 2:** Multivariable linear regression for the difference in numbers of visually fixated vital signs between the technologies. Clustering of observations within the same participant was addressed using cluster robust standard errors. Only the scenario duration significantly affected the difference between technologies.

| Variable | Difference | 95% confidence interval | Standard error | Difference / standard error | P-Value |
|---|---|---|---|---|---|
| Scenario duration (3 vs. 10 seconds) | -1.46 | -2.84 to -0.07 | 0.68 | -2.15 | 0.04 |
| Profession | 0.33 | -0.87 to 1.52 | 0.59 | 0.57 | 0.58 |
| Study center | -0.41 | -1.95 to 1.14 | 0.76 | -0.54 | 0.60 |
| Gender | 0.01 | -1.20 to 1.22 | 0.59 | 0.02 | 0.99 |
| Order of scenarios | 0.10 | -1.51 to 1.70 | 0.78 | 0.12 | 0.9 |
| Technology (conventional vs. avatar [intercept]) | -3.28 | -3.86 to -2.69 | 0.29 | -11.47 | <0.001 |

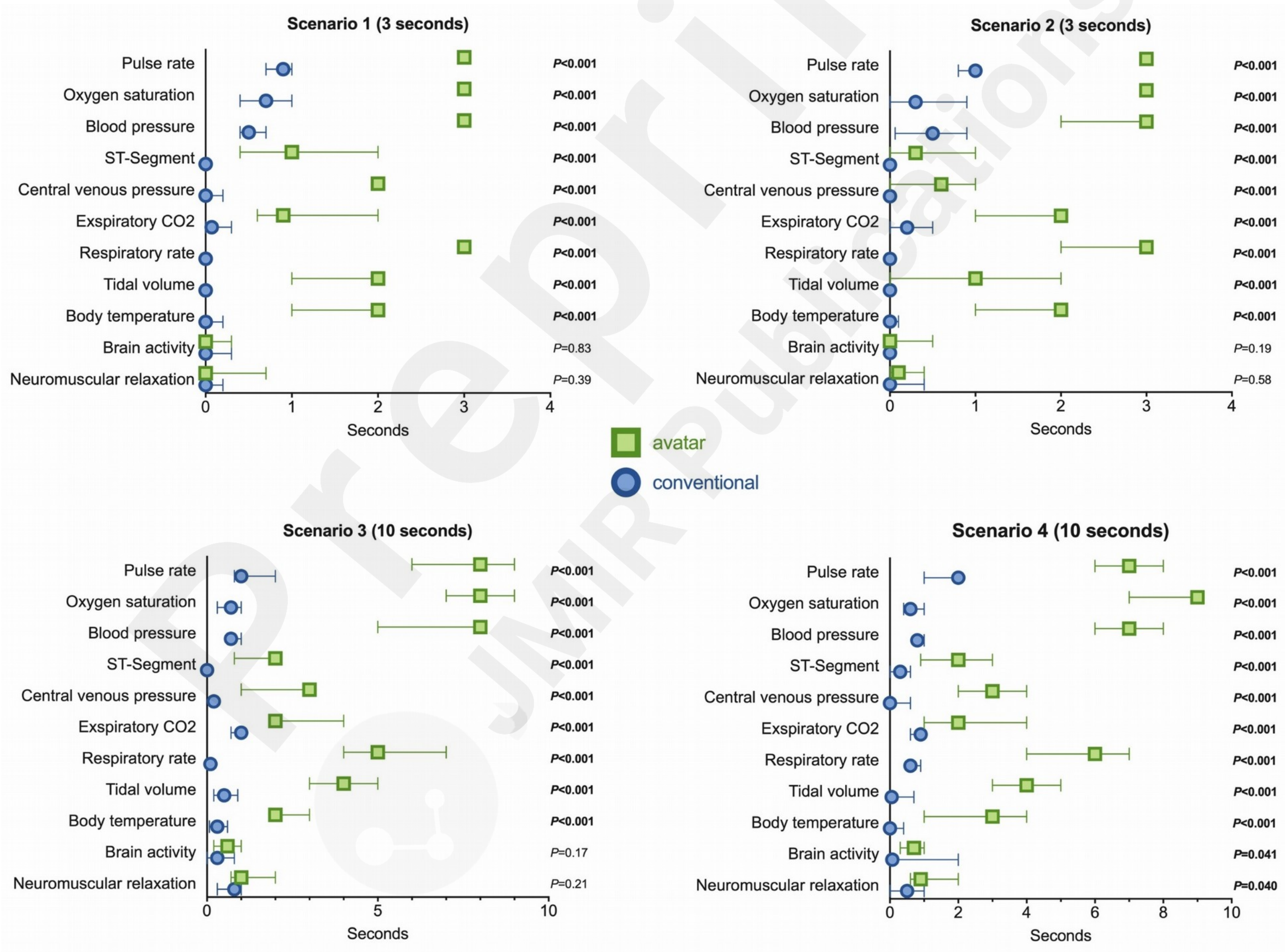
## Visual fixations per vital sign

The analyses for each vital sign individually (Supplementary Figure 2) showed that with avatar-based monitoring nine of the 11 total vital signs were fixated statistically significantly more often per scenario than with conventional patient monitoring. The vital signs pulse rate, oxygen saturation and blood pressure were visible in almost every fixation of participants in all four scenarios with avatar-based monitoring. In comparison, with conventional monitoring, each vital sign was readable only during a small number of visual fixations per observation.

*Duration of visual fixations per vital sign*

As outlined in Figure 2, similar to the number of fixations per vital sign, in all four scenarios, nine of 11 vital signs were statistically significantly longer visually fixated with the avatar than with conventional patient monitoring. With avatar-based monitoring, four critical vital signs: 1. pulse rate, 2. blood pressure, 3. oxygen saturation, and 4. respiratory rate were visible to users during almost the entire time of the scenarios. This was because in the design of the avatar this information is displayed in the form of large anatomical objects which extend across large parts of the screen.

**Figure 2:** Avatar-based monitoring compared to conventional patient monitoring: Median (with interquartile range) durations of visual fixations for each vital sign, scenario, and technology. For example, in scenario 1 (3sec) the median duration of the visual fixations of the participants within 2 cm or less of the vital sign pulse rate was 1 second. In the same scenario with avatar-based monitoring the median duration of the visual fixations of the participants within 2 cm or less of the vital sign pulse rate was 3 seconds. N for scenario 1 (3sec) was 12, for scenario 2 (10sec) 15, for scenarios 3 (3sec) 14, and for scenario 4 (10sec) 15.

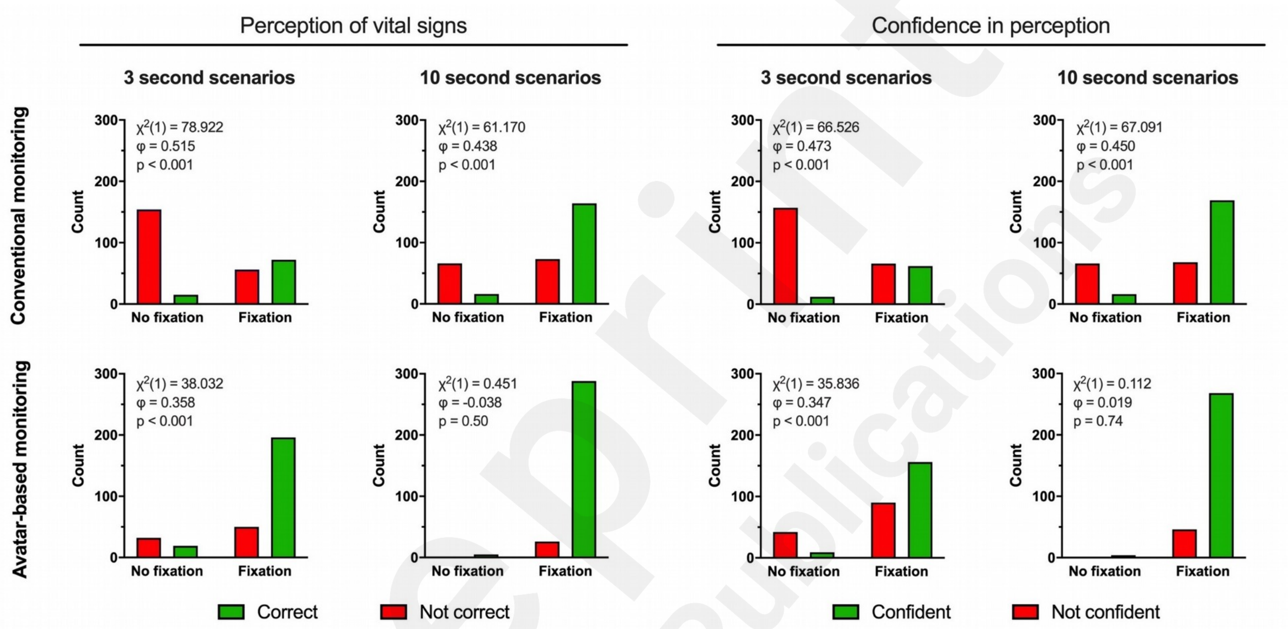


## Correlation of vital sign visual fixations with correct perception and perceived confidence:

A chi-square test for association was conducted between visual fixation of a vital sign and the correct perception of said vital sign. All expected cell frequencies were greater than five, except in the ten second scenario with avatar-based monitoring. In this scenario most participants were able to fixate on every vital sign and perceive it correctly. Accordingly, there was a statistically significant association between visual fixation of a vital sign and the correct perception of said vital sign in the three and ten second scenarios with conventional monitoring, and in the three second scenario with the avatar-based monitoring. When significant, the association was moderately strong (Figure 3).

Similar results were achieved by a chi-square test for association between visual fixation of a vital sign and the participants confidence in having it perceived correctly (Figure 3).

**Figure 3:** Cross tabulation bar graphs with chi-square tests for association between visual fixation of a vital sign and the correct perception of said vital sign, as well as the confidence in the correct perception. Upper row is for conventional monitoring and bottom row is for avatar-based monitoring.



## Discussion

Patient monitoring is a central part of modern surgery, anesthesia, and intensive care.[1,22] While currently available monitors have been enhancing perioperative safety,[23,24] they still mainly show vital sign information as numbers and curve forms, which is not an ideal format for quick and easy information transfer to human users.[4,5] However, an alternative monitoring technique, using an avatar-based representation of the vital sign situation, has been found to enable anesthesiologists to grasp more vital sign information in a shorter time with improved diagnostic confidence and diminished perceived workload.[8,9,21]

In this study, we evaluated eye-tracking data collected in two groups of anesthesiologists at two study centers. We recorded these data while the anesthesiologists were given the task to perceive vital sign information from patient monitoring scenarios presented in the two technologies, i.e., conventional and avatar-based. Specifically, we evaluated how many vital signs and for how long these vital signs could have potentially been read by the participants according to the paths of their foveal or sharp vision across the screen. We found that participants were able to visually fixate more vital signs during the same time with avatar-based monitoring than with conventional patient monitoring. Nine of the 11 total vital signs were fixated more frequently per observation with avatar-based monitoring. Moreover, with avatar-based monitoring, participants fixated the vital signs for longer time intervals per recording, which might give them more time to process the information. More time to perceive the information in turn may have been responsible for the reductions in perceived workload. In short, with the avatar, users see more information for a longer time. These findings were a consequence of the avatars design with many of the vital signs spread out over a large part of the screen and some visualized multiple times. For example, the vital sign “respiratory rate” can be interpreted by looking at the expiratory carbon dioxide “cloud” of the avatar as well as in the excursions of its lungs. The number of correctly perceived vital signs without a visual fixation accounted for less than 10% of the correctly perceived vital signs in all scenarios and with both technologies. This may have been influenced by the audio signal which was played in the conventional monitoring scenarios and which contained information on pulse rate and oxygen saturation. There may also have been some correct guesses without actual perception. In the avatar scenarios, some of the vital signs may have been perceived by the use of peripheral vision, which we found to be an additional advantage of avatar-based monitoring.[21]

The cases of visually fixated vital signs that were not correctly detected accounted for between 0 and 20% of all vital signs, depending on the scenario and technology and may be explained by losses during processing after visual fixation, for example, when a vital sign is forgotten or being confused for another vital sign before being recalled. Numbers are glyphs, which cannot be attributed solely to one vital sign, i.e., it might be possible that a participant, although remembering the value of a number correctly, may miss-attribute the number to another vital sign with a similar range. Indeed, our data showed that when participants had to remember more than just a few vital signs in the more extended 10-second scenarios, the number of vital signs with a visual fixation that participants could not recall was more than twice as high in conventional monitoring than in avatar-based monitoring. These results correlate with research on the holding capacity of our working memory, which has shown that people can only store seven plus minus two digits in short term memory.[25]

With avatar-based monitoring, nearly all participants were able to visually fixate and correctly perceive nearly all of the vital sign information in the longer (10 sec) scenarios. Here, this study shows the limitations of the single-sensor single-indicator design of conventional patient monitoring where a single sensor on the patient feeds a single indicator on the patient monitor. The numbers must be individually read one after the other and afterwards interpreted before meaning and subsequently a global mental picture of the situation can be derived.[5,15] With avatar-based monitoring, we found that the four critically important vital signs pulse rate, blood pressure, oxygen saturation and respiratory rate remained perceptible during almost the entire duration of the monitoring scenarios. To perceive this same information, conventional monitoring requires four visual fixations, the eye-movements in between them and brainwork to interpret the meaning of the values. Avatar-based monitoring facilitates the interpretation work by the use of vital sign visualizations that have a logical commonality with the real phenomena they mirror and therefore do not require further mental translation by the user to be understood. The principle that a good model reflects the reality it represents is found both in principles of logic[14] and in situation awareness design principles, where it is known as "presenting information directly."[4] Anesthesiologists have mentioned information overload as a common problem in connection with patient monitoring.[26] As more and more inexperienced users will likely monitor patients in the future, ease of information transfer will be of paramount importance.[27] The ultimate benefit of the Visual Patient should be an increase in patient safety. Although at this stage of its development, we are not yet able to evaluate patient outcomes, the results of this and our previous studies fit into the context of situation awareness, decision making, and performance. Care providers must perceive and understand the available information before they can confidently make the correct decision and take the right measure.[3,4,28] Situation awareness failures have been identified as root causes of critical anesthesia-events.[29,30]

### Limitations

This study has some important limitations. For one, self-enrollment, based on interest in the technology, could have led to a selection bias. Less technology-savvy care providers may have achieved different results. Secondly, we recorded the source data in a simulated environment. Since the operating room and intensive care unit environment is very complex in real life, it is impossible to predict precisely how substantial the effects of an avatar-based monitor would be in these settings. However, it is plausible that effects would persist when used in a real patient monitor, as the general physiological specifications of information intake do not change. A study with a high level of realism patient simulator in a realistic environment with the technology must be carried out as the next step of scientific evaluation on the way to a commercial product. Another potential limitation is the versatility of the eye tracking method. Although we were able to validate the method through the positive correlation between visual fixation and correct perception, there are influences in perception that are not fully detectable by eye-tracking, such as the influences of the audio signal, peripheral vision, and working memory. Particular strengths of this study include its multi-center design and the balanced enrollment of the different occupational groups and genders — a multivariate regression analysis rendered significant local effects, gender, profession, and scenario ordering effects unlikely. The within-subject study design minimizes the impact of inter-viewer variability of the eye-tracking method and other inter-participant variabilities. The sample size adequately powered the analyses, and the significant magnitude and consistency of the differences observed between the two monitoring technologies increase the internal validity of the study.

### Conclusions

This study analyzed eye-tracking data to explain the improved information transfer with avatar-based patient monitoring. The avatar's design, in which the vital sign information is presented as large scale, integrated, colorful, and direct visualizations allows users to see information about more vital signs with every glance and also see the vital sign information for a longer time with every glance at the monitor. In short, the way the avatar presents the information enables parallel perception of multiple vital signs at the same time, thereby increasing the number of visually fixated vital signs and the time available to view each vital sign. This study provides important groundwork for the future clinical validation of the concept and future studies should examine the technology's performance in simulator-based and subsequently real-life studies.

Multimedia Appendix 1: [Supplementary Video 1, Visual Patient educational video]

Multimedia Appendix 2: [Supplementary Video 2, scenario and eye tracking methodology examples]

## References

1. World Health Organization. Guidelines for Safe Surgery 2009: Safe Surgery Saves Lives. Geneva, Switzerland; 2009.
2. Weiser TG, Haynes AB, Molina G, Lipsitz SR, Esquivel MM, Uribe-Leitz T, Fu R, Azad T, Chao TE, Berry WR, Gawande AA. Size and distribution of the global volume of surgery in 2012. Bull World Health Organ. 2016 Mar 1;94(3):201-9F. PMID: 26966331.
3. Schulz CM, Endsley MR, Kochs EF, Gelb AW, Wagner KJ. Situation awareness in anesthesia: concept and research. Anesthesiology. 2013 Mar;118(3):729-42. PMID: 23291626.
4. Endsley M, Jones, D. Designing for situation awareness: an approach to user-centered design. Boca Raton, FL, USA: CRC Press Inc.; 2011.
5. Drews FA, Westenskow DR. The Right Picture Is Worth a Thousand Numbers: Data Displays in Anesthesia. Human Factors. 2006;48(1):59-71. PMID: 16696257.
6. Ford S, Birmingham E, King A, Lim J, Ansermino JM. At-a-glance monitoring: covert observations of anesthesiologists in the operating room. Anesth Analg. 2010 Sep;111(3):653-8. PMID: 20581165.
7. Loeb RG. Monitor surveillance and vigilance of anesthesia residents. Anesthesiology 1994;80:527-33. PMID: 8141449.
8. Tscholl DW, Handschin L, Neubauer P, Weiss M, Seifert B, Spahn DR, Noethiger CB. Using an animated patient avatar to improve perception of vital sign information by anaesthesia professionals. Br J Anaesth. 2018 Sep;121(3):662-71. PMID: 30115265.
9. Tscholl DW, Weiss M, Handschin L, Spahn DR, Nothiger CB. User perceptions of avatar-based patient monitoring: a mixed qualitative and quantitative study. BMC Anesthesiol. 2018 Dec 11;18(1):188. PMID: 30537934.
10. Traquair HM. An introduction to clinical perimetry. London: Kimpton; 1927.
11. Holmqvist K. Eye tracking: a comprehensive guide to methods and measures. Oxford, United Kingdom: Oxford University Press; 2011.
12. J Nielsen KP. Eyetracking Web Usability. Berkeley, CA, USA: New Riders; 2009.
13. Rayner K. Eye movements in reading and information processing: 20 years of research.

Psychol Bull. 1998 Nov;124(3):372-422. PMID: 9849112.

14. Wittgenstein L. Tractatus logico-philosophicus. London, United Kingdom: Routledge & Kegan Paul; 1922.
15. Degani Asaf JC, Iverson David, Shafto Michael, Olson Leonard. On Organization of Information: Approach and Early Work. Moffett Field, CA, USA; 2009.
16. Endsley M, Garland, D. Situation Awareness Analysis and Measurement. Boca Raton: CRC Press; 2000.
17. Hart SG, Stavenland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. Human Mental Workload: Elsevier; 1988. p. 139-83.
18. Hart S. Nasa-Task Load Index (Nasa-TLX); 20 Years Later. Human Factors and Ergonomics Society Annual Meeting; 2006.
19. Tscholl DW, Weiss M, Spahn DR, Noethiger CB. How to Conduct Multimethod Field Studies in the Operating Room: The iPad Combined With a Survey App as a Valid and Reliable Data Collection Tool. JMIR Res Protoc. 2016 Jan 5;5(1):e4. PMID: 26732090.
20. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007 May;39(2):175-91. PubMed PMID: 17695343.
21. Pfarr J, T Ganter M, Spahn D, Nöthiger C, W Tscholl D. Avatar-based patient monitoring with peripheral vision: A Multicenter Comparative Eye-Tracking Study. J Med Internet Res. 2019 Jul 17;21(7):e13041. PMID: 31317870.
22. Gelb AW, Morriss WW, Johnson W, Merry AF, Abayadeera A, Belii N, et al. World Health Organization-World Federation of Societies of Anaesthesiologists (WHO-WFSA) International Standards for a Safe Practice of Anesthesia. Anesth Analg. 2018 Jun;126(6):2047-55. PMID: 29734240.
23. Moller JT, Johannessen NW, Espersen K, Ravlo O, Pedersen BD, Jensen PF, et al. Randomized evaluation of pulse oximetry in 20,802 patients: II. Perioperative events and postoperative complications. Anesthesiology. 1993 Mar;78(3):445-53. PMID: 8457045
24. Maheshwari K, Khanna S, Bajracharya GR, Makarova N, Riter Q, Raza S, et al. A Randomized Trial of Continuous Noninvasive Blood Pressure Monitoring During Noncardiac Surgery. Anesth Analg. 2018 Aug;127(2):424-31. PMID: 29916861.

25. Miller GA. The magical number seven, plus or minus two: some limits on our capacity for processing information. 1956. Psychol Rev. 1994 Apr;101(2):343-52. PMID: 8022966.
26. Tscholl DW, Handschin L, Rössler J, Weiss M, Spahn DR, Nöthiger CB. It's not you, it's the design - common problems with patient monitoring reported by anesthesiologists: a mixed qualitative and quantitative study. BMC Anesthesiol. 2019 May 28;19(1):87. PMID: 31138143.
27. Michard F, Sessler DI. Ward monitoring 3.0. Br J Anaesth. 2018 Nov;121(5):999-1001. PMID: 30336872.
28. Endsley MR JD. Situation awareness information dominance & information. Wright-Patterson Air Force Base, OH, USA: 1997.
29. Schulz CM, Burden A, Posner KL, Mincer SL, Steadman R, Wagner KJ, et al. Frequency and Type of Situational Awareness Errors Contributing to Death and Brain Damage: A Closed Claims Analysis. Anesthesiology. 2017 Aug;127(2):326-37. PMID: 28459735.
30. Schulz CM, Krautheim V, Hackemann A, Kreuzer M, Kochs EF, Wagner KJ. Situation awareness errors in anesthesia and critical care in 200 cases of a critical incident reporting system. BMC Anesthesiology. 2016 January 16;16(1):4. PMID: 26772179.
31. Engelken EJ, Stevens KW, McQueen WJ, Enderle JD. Eye tracking performance variability in a homogeneous population. Biomed Sci Instrum. 1995;31:229-33. PMID: 7654968.
32. Coey CA, Wallot S, Richardson MJ, Orden GV. On the Structure of Measurement Noise in Eye-Tracking. JEMR [Internet]. 2012Sep.4 [cited 2019Nov.14];5(4). Available from: https://bop.unibe.ch/JEMR/article/view/2343